Apendix A : Tables for the DUI

1.3 The Data Usage Index

The "Data Usage Index" (DUI) developed by Peter Ingwersen and Vishwa Chavan (2011; 2012) includes the following elements :

	Code	Indicator		Explanation
1	s (u)	Searched Records		Number of records searched/viewed (by IP address) in unit
2	d(u)	Download Frequency		Number of downloaded records from unit
3	r(u)	Record Numbers		Number of records in (period; dataset(s); geographical and/or species unit)
4	S(u)	Search Events		Number of different searches (by IP address) in unit
5	D(u)	Download Events		Number of different downloads from unit
6	R(u)	Dataset Number		Number of datasets in (period, geographical and/or species unit)
7	s(u)/ S(u)	Search Density		Average number of searched records per search event
8	d(u) / D(u)	Download Density		Average download frequency per download event
9	d(u) / r(u)	Usage Impact		Download frequency per stored record per unit
10	s(u) / r(u)	Interest Impact		Searched records per stored record per unit
11	d(u) / s(u)	Usage Ratio		Ratio of download frequency to searched records in unit
12	D(u) / S(u)	Usage balance		Ratio of download events to search events for unit (in %)
13	U(u) / r(u)	Usage Score		Ratio of unique downloaded records (U) to record number (in %)
14	I(u) / r(u)	Interest Score		Ratio of unique searched records (I) to record number (in %)

Table 1. The Indicators for Ingwersen and Chavan’s Data Use Index

The DUI as adapted for the RDA

	Code	Indicator	Explanation
1	uu(ds)	Unique Users	Unique users that downloaded data during a time window
2	n(ds)	Number of Datasets	Number of Datasets assigned DS number by RDA
3	f(ds)	Files DS	Number of files in Dataset per time window
4	d(ds)	Download Frequency	Total number of files downloaded per time window
5	hp(ds)	Homepage Hits	Home Page Hits of Data Set per time window
7	d(ds) /uu (ds)	Download Density	Average number of files downloaded per unique user
8	d(ds) / f (ds)	Usage Impact	Total number of downloaded files over total files in dataset
9	d(ds) / hp(ds)	Usage Balance	Files downloaded by number of homepage hits per time window
10	hp(ds) / f(ds)	Interest impact	Total homepage hits per number of files in dataset
11	hp(ds) / uu(ds)	Secondary Interest Impact	Total Homepage hits over unique users
12	ss(ds) / d(ds)	Subset Ratio	Subset requests over total number files downloaded

Table 2: The Indicators for the RDA's Data Use Index

Our pilot project included the analysis of ICOADS (ds540.0); a set of global observational data (ds093.0), and a popularly analyzed data product from a numerical weather prediction center (ds083.2).

Indicator	ds540.0-1 : 3/2011	ds540.0-1 : 7/2012	ds083.2 : 03/2011	ds083.2: 05/2012	ds093.0-3 : 3/2011	ds093.0-6 : 07/2012
Unique Users	46	45	987	976	88	272
Download Frequency	264	373	374962	335422	3528	23739
Files DS	433	473	22221	25504	195616	277642
Homepage Hits	685	588	6749	6907	1655	3534
Subset Requests	145	35	n/a	42	175	791
Download Density	5.73913043	8.28888888	379.900709	343.67008	40.0909090	87.275735
Usage Impact	0.60969976	0.78858351	16.8736485	13.151740	0.01803528	0.0855021
Interest Impact	1.58198614	1.24312896	0.30371145	0.2708202	0.00846043	0.0127286
Download Ratio	2.59469697	1.57640750	0.01799915	0.0205919	0.46910430	0.1488689
Usage Balance	0.38540146	0.63435374	55.5581567	48.562617	2.13172205	6.7173174
Subset Ratio	0.54924242	0.09383378	n/a	0.0001253	0.04960318	0.0333207
Datasets	2	2	1	1	3	6
Secondary Interest Impact	14.8913043	13.0666666	6.83789260	7.0768442	18.8068181	12.992647

Table 3. Indicator scores for ds540; ds083; and ds093.6